

BAB I

PENDAHULUAN

Most Muslims in Indonesia are not used to writing Arabic so writing with Latin queries is needed. The search system called Lafzi¹ is one example of an effective application for searching Arabic queries based on sound similarity. The Lafzi application helps correct writing errors due to the sound of Arabic letters that are almost the same as the pronunciation but not correct errors due to typos. In the other side, Latin writing will usually cause writing errors or typos such as the "Bisnillah" query which should be "Bismillah" (In the name of Allah). Writing errors occur because the position of the letter "N" on the QWERTY type keyboard is adjacent to the letter "M"². This writing error is usually ignored by the users and causes the results of the search to not be found. The system for generating suggestions is needed to correct errors in Latin queries so that the initial query can be corrected and give the desired search results.

In this research, the Query Correction system will correct errors from the queries entered by the user. The Query Correction system using auto-complete to equipped missing trigram and the edit distance metric to calculate the differentiation value between the suggested query with the initial query. The way the system works is to separate and sort trigram tokens from queries (user input) based on the verse. Each verse that has a missing trigram token will be equipped and re-transformed into a corrected query. Each corrected query will be compared to the edit distance value against the initial query (input from the user), then a corrected query will be taken which has the smallest edit distance value and will be made as a suggested query.